Atty. Docket No. DE920000055US1
(590.080)

REMARKS

In the Office Action dated June 1, 2005, pending Claims 19-49 were rejected and the rejection made final. Claims 19, 25, 31, 39, 47-49 are independent claims; the remaining claims are dependent claims. Applicants and the undersigned are most grateful for the time and effort accorded the instant application by the Examiner. The Office is respectfully requested to reconsider the rejections presented in the outstanding Office Action in light of the following remarks.

Claims 25-30, 39-41, 43-47 and 49 stand rejected under 35 USC § 102(b) as being anticipated by Kimber et al. Reconsideration and withdrawal of this rejection is respectfully requested.

As best understood, Kimber et al. appears to be directed to a method of clustering speaker data from a plurality of unknown speakers in conversational data. (Abstract; Col. 1, lines 26-28) While the identify of the speakers is not known in Kimber et al., speech segments from each speaker appear to be clustered together. The data clusters are indexed indicating multiple speakers and audio segments that correspond to different speakers. (column 5, lines 33-45) An index for the complete recording is then created by collecting the various segments that are similarly marked by an individual. (Col. 12, lines 17-19) The outstanding Office Action states that Kimber et al. teach identification of speakers from a plurality of speakers using speech models. However, Kimber et al. also states that the speakers to be indexed and recognized may not even be known before the speech data is processed. (column 8, lines 23-28) Thus, although Kimber et al. may be

- 12 -

Atty. Docket No. DE920000055US1
(590.080)

able to identify a known speaker from a plurality of speakers, it is respectfully submitted

that Kimber et al. may not necessarily have any known speakers in the data stream that is

being processed and thus do not meet the limitations of the invention as claimed above.

Further,. Kimber et al. do not teach indexing or transcribing the audio stream with respect

to the detected speaker change only if the known speaker is recognized. Kimber et al.

teach indexing and transcribing the entire audio stream, rather than portions of the stream

as taught by Applicant.

Creating an index for a complete conversation, or for that matter, unknown

speakers, stands in stark contrast to the present invention. As discussed in the

specification, there are various issues with training speech recognition systems to deal

with multiple speakers. Prior efforts to automize the indexing of audio material, e.g.,

using prior are speech recognition technology, thus failed due to the large variability of

speech styles and dialects of the human individuals engaged in those interactions. (Page

2, lines 8-10) Thus, the idea underlying the present invention is to locate segments in a

continuous audio stream where a change-over to at least one predefined speaker occurs

and to apply speech recognition or voice control techniques only to those audio segments

belonging to the predefined speakers. (Page 3, lines 13-16) In particular, the present

invention proposes to apply known speaker recognition techniques to conversations

between a well-known speaker and a multitude of unknown speakers and thereby allows

to transcribe only the utterances of the well-known speaker as an index and summary of

the dialogues. (Page 4, lines 7-10) However, Kimber et al. apply speaker recognition

techniques and indexing to all speakers in the conversation and create an index for each

- 13 -

speaker in the conversation. As stated above, the speakers to be indexed and recognized may not even be known before the speech data is processed. (column 8, lines 23-28)

Independent Claim 19 recites, *inter alia*, <u>identifying a known speaker from among the plurality of speakers</u> and <u>transcribing at least part of the continuous audio stream if the known speaker is recognized</u>. (emphasis added) Independent Claim 25 recites, *inter alia*, <u>indexing the audio stream with respect to the detected speaker change if the known speaker is recognized</u>. (emphasis added) Similar language also appears in the other Independent Claims.

It is respectfully submitted that Kimber et al. clearly falls short of present invention (as defined by the independent claims) in that, *inter alia*, it does not disclose identifying a known speaker from among the plurality of speakers *and* transcribing or indexing at least part of the continuous audio stream if the known speaker is recognized (emphasis added) Accordingly, Applicants respectfully submit that the applied art does not anticipate the present invention because, at the very least, "[a]nticipation requires the disclosure in a single prior art reference of each element of the claim under construction." *W.L. Gore & Associates, Inc. v. Garlock*, 721 F.2d 1540, 1554 (Fed. Cir. 1983); *see also In re Marshall*, 198 U.S.P.Q. 344, 346 (C.C.P.A. 1978).

Claims 19-24, 31-38, 42 and 48 stand rejected under 35 USC § 103(a) as obvious over Kimber et al. in view of Glickman et al. Specifically the Office asserted that "[i]t would have been obvious ... to modify Kimber et al. by incorporating the teaching of Glickman et al. in order to provide automatic closed-caption using speaker-dependent

- 14 -

Atty. Docket No. DE920000055US1
(590.080)

models to enhance speech recognition accuracy." Reconsideration and withdrawal of the present rejections are hereby respectfully requested.

A 35 USC 103(a) rejection requires that the combined cited references provide both the motivation to combine the references and an expectation of success. Not only is there no motivation to combine the references, no expectation of success, but actually combining the references would not produce the claimed invention. Thus, the claimed invention is patentable over the combined references and the state of the art.

Glickman et al. does not overcome the deficiencies of Kimber et al. set forth above. In that regard, Glickman et al. "develop[s] separate acoustic-phonetic models ... for [each of] the multiple speakers." (Col. 5, lines 51-52) Glickman et al. continues that "[a]fter the models are 'trained', speaker recognition can automatically be performed" and "[t]his technique can also be used to perform 'automatic' closed captioning." (Col. 5, lines 52-55) Glickman et al. thus teaches transcription of the speech of all speakers involved in a particular conversation. Further, Glickman et al. teach storing only those words that are recognized in the audio stream (column 3, lines 36-44). This is in stark contrast to the present invention, in which all of the words in the audio stream are indexed or transcribed that correspond to a known speaker.

There is an inherent tension in Kimber et al. and Glickman et al. given that in Kimber et al. speakers may be unknown and in Glickman et al. all of the speakers are necessarily known, which teaches away from combining these two references. At best, however, combining Kimber et al. and Glickman et al. would result in continuously
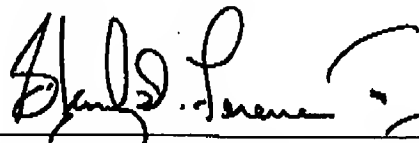
- 15 -

training models for each speaker so that the recognized speech of all participants in a conversation would be transcribed and displayed as closed captioning. Even if there were a motivation for the combination, this combination does not teach or suggest the claimed invention.

In view of the foregoing, it is respectfully submitted that Independent Claims 19, 25, 31, 39 and 47-49 fully distinguish over the applied art and are thus allowable. By virtue of dependence from Claims 19, 25, 31 and 39, it is thus also submitted that Claims 20-24, 26-30, 32-38 and 40-46 are also allowable at this juncture.

In summary, it is respectfully submitted that the instant application, including Claims 19-49, is presently in condition for allowance. Notice to the effect is hereby earnestly solicited. If there are any further issues in this application, the Examiner is invited to contact the undersigned at the telephone number listed below.

Respectfully submitted,

Stanley D. Ference III
Registration No. 33,879

Customer No. 35195
FERENCE & ASSOCIATES
409 Broad Street
Pittsburgh, Pennsylvania 15143
(412) 741-8400
(412) 741-9292 - Facsimile

Attorneys for Applicants

- 16 -